Purves

methods for area-under-the-curve (*AUC*)
J. Pharmacokin. Biopharm. 20:211-226

statistical moment theory to the evaluation
ne. *J. Pharmacokin. Biopharm.* 8:509-534

harmacokinetics. Definition, computation
acokin. 17:424-440 (1989).
Analysis, Allyn and Bacon, Boston, 1988.
ependent random variables. *Am. Statist.*

egies for noncompartmental estimation of
n. 11:435-446 (1983).
and W. T. Vetterling. *Numerical Recipes:*
niversity Press, Cambridge, 1985.

Journal of Pharmacokinetics and Biopharmaceutics, Vol. 20, No. 5, 1992

PHARMACOMETRICS

Building Population Pharmacokinetic–Pharmacodynamic Models. I. Models for Covariate Effects

Jaap W. Mandema,[1] Davide Verotta,[1] and Lewis B. Sheiner[1,2,3]

Received June 8, 1992—Final September 10, 1992

One major task in clinical pharmacology is to determine the pharmacokinetic–pharmacodynamic (PK-PD) parameters of a drug in a patient population. NONMEM is a program commonly used to build population PK-PD models, that is, models that characterize the relationship between a patient's PK-PD parameters and other patient specific covariates such as the patient's (patho)physiological condition, concomitant drug therapy, etc. This paper extends a previously described approach to efficiently find the relationships between the PK-PD parameters and covariates. In a first step, individual estimates of the PK-PD parameters are obtained as empirical Bayes estimates, based on a prior NONMEM fit using no covariates. In a second step, the individual PK-PD parameter estimates are regressed on the covariates using a generalized additive model. In a third and final step, NONMEM is used to optimize and finalize the population model. Four real-data examples are used to demonstrate the effectiveness of the approach. The examples show that the generalized additive model for the individual parameter estimates is a good initial guess for the NONMEM population model. In all four examples, the approach successfully selects the most important covariates and their functional representation. The great advantage of this approach is speed. The time required to derive a population model is markedly reduced because the number of necessary NONMEM runs is reduced. Furthermore, the approach provides a nice graphical representation of the relationships between the PK-PD parameters and covariates.

KEY WORDS: pharmacokinetics; population analysis; model building; generalized additive models; NONMEM.

This work was supported in part by U.S. Department of Health and Human Services grants GM 26676 and GM 26691. Jaap W. Mandema was supported by a NATO Science Fellowship awarded by the Netherlands Organization for Scientific Research.

[1]Department of Pharmacy, School of Pharmacy, University of California, San Francisco, California 94143.

[2]Department of Laboratory Medicine, School of Medicine, University of California, San Francisco, California 94143.

[3]To whom correspondence should be addressed at the Department of Laboratory Medicine, School of Medicine, Box 0626, University of California, San Francisco, California 94143.

0090-466X/92/1000-0511$06.50/0

INTRODUCTION

Studies of population pharmacokinetics and pharmacodynamics have led to the appreciation of the great degree of variability in pharmacokinetic–pharmacodynamic (PK-PD) parameters across patients. Numerous studies have quantified the effects of factors, such as age, gender, disease states, concomitant drug therapy, etc., on the pharmacokinetics and pharmacodynamics of drugs, with the goal of accounting for interindividual variability. NONMEM is a widely used program for pharmacokinetic and pharmacodynamic population analysis (1). NONMEM describes interindividual variability in terms of fixed and random effects. The fixed effects relate PK-PD parameters to covariates. The interindividual random effects quantify the residual unexplained variability. An additional random effect quantifies intraindividual and measurement variability.

Finding a population model that adequately describes the data can be a painstaking and complicated task. This task comprises not only finding the covariates that significantly influence the PK-PD parameters but also determining the shape of the relationships between covariates and parameters. In particular when the PK-PD models become complicated and the effects of numerous covariates must be considered, the number of possible models increases dramatically. Obviously, one cannot try all possible models because the computer time required would be unacceptable.

In this paper we describe an extension of a previously suggested approach to efficiently find the relationships between covariates and PK-PD parameters (2). Four real-data examples are used to illustrate the stepwise procedure to building a population model that adquately describes the data. In describing the approach, we have divided the paper in several sections. The general procedure of the analysis is described in the first section. In the Data section, the real-data examples are outlined; the Application section gives the application of the analysis to the real-data examples, followed by a short Discussion section.

GENERAL PROCEDURE OF NEW APPROACH

The data analysis is performed using a stepwise approach. Briefly, individual empirical Bayes estimates of PK parameters are obtained. Subsequently, the relationship between the individual PK parameter estimates and covariates is modeled using a generalized additive model (GAM) to allow nonlinear covariate-parameter relationships to be discovered. Finally, the population model is built using NONMEM on the basis of the GAM analysis of the previous step. The procedure is now described in greater detail.

tics and pharmacodynamics have
of variability in pharmacokinetic–
across patients. Numerous studies
ch as age, gender, disease states,
rmacokinetics and pharmacodyn-
ing for interindividual variability.
harmacokinetic and pharmacody-
M describes interindividual vari-
ts. The fixed effects relate PK-PD
dual random effects quantify the
ditional random effect quantifies
ity.
equately describes the data can be
s task comprises not only finding
the PK-PD parameters but also
s between covariates and param-
dels become complicated and the
onsidered, the number of possible
one cannot try all possible models
d be unacceptable.
nsion of a previously suggested
ps between covariates and PK-PD
are used to illustrate the stepwise
that adquately describes the data.
ded the paper in several sections.
escribed in the first section. In the
outlined; the Application section
e real-data examples, followed by

PPROACH

a stepwise approach. Briefly, indi-
parameters are obtained. Subse-
idual PK parameter estimates and
additive model (GAM) to allow
ps to be discovered. Finally, the
M on the basis of the GAM analy-
now described in greater detail.

1. In a first step, a basic population PK model without any covariates is estimated. In all examples the NONMEM (Version 4, level 1.0) model for the jth concentration measurement in the ith individual is given by

$$C_{ij}=f(\mathbf{p}_j, t_{ij})(1+\varepsilon_{ij}) \quad (1)$$

where f is the model predicting the jth concentration (usually a mono- or biexponential) and $\mathbf{p}_i=(p_{1i}, p_{2i}, \ldots, p_{mi})$ are the m PK parameters of the ith individual. ε_{ij} represents the residual departure of the model from the jth observation available from individual i. The kth element of $\mathbf{p}_i$ is modeled as

$$p_{ki}=\theta_k \exp(\eta_{ki}) \quad (2)$$

where θ_k is the population mean of p_{ki} and $\exp(\eta_{ki})$ expresses the (random) difference between θ_k and p_{ki}. $\{\varepsilon_{ij}\}$ are assumed independent normally distributed with mean zero and variance σ^2. $\eta_i=(\eta_{1i}, \eta_{2i}, \ldots, \eta_{mi})$ are assumed to be independent multivariate normally distributed, with mean zero and with variance-covariance the diagonal matrix Ω with diagonal elements $(\omega_1^2, \ldots, \omega_m^2)$. We assume no covariance between elements of η at this stage to be sure that each covariate has the opportunity to appear to be related to each η. Subsequently, empirical Bayes estimates of the p_{ki} are obtained by standard methods using the estimated values of θ_k, σ^2 and Ω (3,4).

2. With the estimates of p_{ki} from Step 1 treated as "data," this step corresponds to the now classical regression problem of variable selection, and we take advantage of the considerable recent work done on this problem by others (e.g., 5). In this step, a regression model is derived to model the dependence of the individual PK parameter estimates p_{ki} on the covariates $X_{1i}, X_{2i}, \ldots, X_{ni}$.

$$p_{ki}=g_k(X_{1i}, \ldots, X_{ni}) \quad (3)$$

where g_k is a multidimensional smooth function. Without some constraints, such a general description of the data would entail serious problems. First, there is the problem of interpreting the function in dimensions higher than 2. Second, there is the problem, as Bellman (6) put it, of the "curse of dimensionality": The number of points required to fill a space to a given density grows exponentially with the dimension of the space but usually the data do not. This makes the estimation of complex multivariate functions using sparse noisy data highly inaccurate. An often-used simplification is to characterize the relationship between parameters and covariates using a multiple linear regression model

$$p_{ki}=\alpha_{k0}+\sum_{l=1}^{n}\alpha_{kl}X_{li} \quad (4)$$

However, this model makes a strong assumption about the dependence of p_k on X_l, namely, that the dependence is linear in each of the predictors.

A more general recent approach is the use of so-called generalized additive models (GAM) (5).

$$p_{ki} = \alpha_{k0} + \sum_{l=1}^{n} g_{kl}(X_{li}) \tag{5}$$

where $g_{kl}(X_{li})$ is an arbitrary univariate function with $\sum_{i=1}^{N_i} g_{kl}(X_{li}) = 0$. The functions $g_{kl}(X_{li})$ can be represented by any function. The use of spline functions (for description of splines see Appendix) is very appealing, because they are flexible, yet easy to use. In a subsequent step the spline functions can be replaced by a parametric representation, if appropriate. By assuming an additive structure, GAM allow straightforward interpretation and display of the role of the various covariates. The price one pays for this simplification is that interactions between covariates are not included. Nonetheless, a GAM is often a useful approximation to the necessarily more complex true multivariate regression surface.

The GAM is built using a stepwise addition/deletion method (7). This method steps through a series of models along a path determined as follows: Each of the covariates is allowed to be left out of the model at each step, or to enter the model in any of several prespecified functional representations. At each step the model is changed by addition (deletion) of the single term that results in the largest decrease in the Akaike information criterion (*AIC*) (8). In this context, the *AIC* is proportional to the residual sum of squares from the GAM fit, but adds a penalty, proportional to the number of parameters in the model. The search is stopped when the *AIC* has reached a minimum value. Calculations were performed using the statistical program S-plus (Version 3.0, Statistical Sciences Inc., Seattle, WA). A Fortran program for GAM fits is also available elsewhere (ref. 5, p. 307). This step represents our elaboration on the method of Maitre *et al.* (2) who simply advise examination of graphical displays at this step so as to proceed to the next and last step.

3. In a third and final step a parametric or semiparametric mixed effects model describing the relationship between covariates and PK parameters is built using NONMEM. The regression models found in the second step serve as an initial guess for the final population model. In the NONMEM analysis the addition of covariance terms between the PK parameters is explored. Plots of the individual PK parameter estimates from Step 1 versus each other serve to indicate which covariance terms should be included.

DATA

Data on four drugs are discussed. They are disguised for confidentiality. The analyses reported here are not meant to be definitive but are presented

ıe of so-called generalized additive

$g_{kl}(X_{li})$ (5)

function with $\sum_{i=1}^{N_l} g_{kl}(X_{li}) = 0$. The
/ any function. The use of spline
ppendix) is very appealing, because
bsequent step the spline functions
tation, if appropriate. By assuming
tforward interpretation and display
rice one pays for this simplification
not included. Nonetheless, a GAM
cessarily more complex true multi-

ıddition/deletion method (7). This
ılong a path determined as follows:
ft out of the model at each step, or
pecified functional representations.
dition (deletion) of the single term
ıkaike information criterion (*AIC*)
ınal to the residual sum of squares
ɔportional to the number of param-
ɔed when the *AIC* has reached a
rmed using the statistical program
nc., Seattle, WA). A Fortran pro-
ewhere (ref. 5, p. 307). This step
d of Maitre *et al.* (2) who simply
at this step so as to proceed to the

ric or semiparametric mixed effects
n covariates and PK parameters is
models found in the second step
ɔulation model. In the NONMEM
ns between the PK parameters is
meter estimates from Step 1 versus
iance terms should be included.

ey are disguised for confidentiality.
t to be definitive but are presented

merely to illustrate the use of the model-building method described above.

Drug A is a nonsteroidal anti-inflammatory that was administered as a single oral dose of 5 mg/kg or 10 mg/kg to 93 children with febrile illness. Two to six plasma samples were obtained from each individual up to 8 hr postdosing (total of 411 plasma samples). The following demographic data were recorded: gender (*SEX*), male (53), female (40); race (*RACE*), Caucasian (17), black (76); location (*LOC*), clinic (65), hospital/inpatient (28); food since dosing (*FED*), no (25), yes (68); dose level (*DRG*), 5 mg/kg (45), 10 mg/kg (48); height (*HT*), 64–148 cm; weight (*WT*), 5.6–45.5 kg; age (*AGE*), 3–132 months.

Drug B is an antiarrhythmic, administered orally to 136 hospitalized men for various arrhythmias. Plasma samples were obtained for routine clinical purposes (total of 361 samples). The following demographic data were recorded: race (*RACE*), Caucasian (91), Latin (35), black (10); smoking (*SMO*), no (91), yes (45); ethanol abuse (*ET*), no ethanol abuse or social drinker (90), current ethanol abuse (16), history of ethanol abuse (30); congestive heart failure (*HF*), no or mild (56), moderate (40), severe (40); creatinine clearance (*RF*), < 50 ml/min (48), >50 ml/min (88); weight (*WT*), 41–119 kg; height (*HT*), 154–202 cm; age (*AGE*), 42–92 years; α-1-acid glycoprotein concentration (*GLP*), 39–316 mg/dl.

Drug C is an antihypertensive, administered twice daily in a dose range of 1 mg/day up to 20 mg/day in patients with mild, moderate, or severe hypertension. Each patient started with the lowest dose and the dose was increased until the patient's arterial pressure was controlled. After at least a period of 8 weeks continuous therapy the pharmacokinetics of Drug C were studied in 64 patients during a 12-hr dosing interval. After 4 weeks, 35 patients were restudied. A total number of 887 plasma samples were obtained. The following demographic data were recorded: gender (*SEX*), male (63), female (36); race (*RACE*), Caucasian (67), black (32); visit number (*VIS*), visit 15 (53), visit 16 (46); Smoking (*TOB*), no (74), yes (25); prior therapy (*PT*), no (3), yes (96); cotreatment with hydrochlorothiazide (*HCTZ*), no (44), yes (55); cotreatment with propranolol (*PROP*), no (85), yes (14); other concomitant therapy (*CON*), no (24), yes (75); food (*FF*), fed (50), fasting (49); height (*HT*), 140–188 cm; weight (*WT*), 51–139 kg; age (*AGE*), 24–69 years; serum creatinine (*SECR*), 0.6–1.8 mg/dl.

Drug D is a broad spectrum antibiotic. The pharmacokinetics of drug D were studied in 74 critically ill patients with infections with organisms sensitive to drug D. The drug was administered twice daily, 200 mg or 400 mg, in 1-hr infusions. For each individual three plasma samples were obtained during a 12-hr dosing interval. In total 113 dosing intervals were studied (337 plasma concentrations). The following demographic data were recorded: age (*AGE*), 18–84 years; weight (*WT*), 43–125 kg; creatinine

clearance (*CLCR*), 0.4–312 mL/min; Glasgow score (*GLAS*), 3–15; simplified acute physiology score (*SAPS*), 1–26; albumin (*ALB*), 17–40 g/L; bilirubin (*BIL*), 4–150 μmol/L; alanine amino transferase (*ALAT*), 3–200 IU/L; alkaline phosphatase (*AP*), 32–615 IU/L; prothrombin level (*PT*), 36–100% normal; systolic blood pressure (*SPB*), 60–175 mmHg; heart rate (*HR*), 60–170 beats/min; artificial ventilation (*AV*), no (42), yes (71); gender (*SEX*), male (86), female (27); center where study was performed (*CEN*), center 1 (44), center 2 (69); administration of high dose catecholamine (*CAT*), no (92), yes (21).

APPLICATION

The stepwise procedure of the new approach is demonstrated in full using drug A. The other examples are used to address certain specific issues. The purpose of the first step is to determine the basic PK model and to obtain empirical Bayes estimates of the model parameters for each individual using NONMEM. A simple one-compartment model with first-order absorption best described the data of Drug A.

Changing Covariates Within One Individual

A complication may arise with the next step, the empirical Bayes estimation of the model parameters, when the value of a covariate changes in time within one individual. Actually, what one wants is an estimate of the model parameters for each constant value of a covariate within one individual. However, some covariates may change literally from one measurement to the next, and one cannot obtain precise empirical Bayes estimates of the model parameters for each separate measurement. To solve this problem each subject's data is subdivided into disjoint contiguous time periods in each of which the covariates are fairly constant. The data corresponding to each of these time periods are considered to be a new "individual." The length of these time periods depends on the number of samples taken, the variability of the covariates, and complexity of the PK model. In this way it is possible to obtain empirical Bayes estimates of period-specific model parameters which can be correlated to the corresponding time averaged values of the covariates during the period. Sometimes this subdivision comes naturally, e.g., when the subjects are studied during different dosing intervals and several samples are taken during each dosing interval. This is the case for Drugs C and D. For those drugs, in fact, some of the covariates changed markedly between the dosing intervals studied, but were constant within the short period of the dosing interval. For each dosing interval, on average, 9 (Drug C) and 3 (Drug D) concentration measurements were available,

lasgow score (*GLAS*), 3–15; simpli-
6; albumin (*ALB*), 17–40 g/L; bili-
amino transferase (*ALAT*), 3–
32–615 IU/L; prothrombin level
ressure (*SPB*), 60–175 mmHg; heart
ventilation (*AV*), no (42), yes (71);
center where study was performed
ministration of high dose catechol-

v approach is demonstrated in full
ed to address certain specific issues.
ermine the basic PK model and to
iodel parameters for each individual
npartment model with first-order
rug A.

lual

ext step, the empirical Bayes estima-
value of a covariate changes in time
e wants is an estimate of the model
a covariate within one individual.
literally from one measurement to
e empirical Bayes estimates of the
easurement. To solve this problem
lisjoint contiguous time periods in
onstant. The data corresponding to
red to be a new "individual." The
the number of samples taken, the
exity of the PK model. In this way
estimates of period-specific model
the corresponding time averaged
1. Sometimes this subdivision comes
lied during different dosing intervals
ch dosing interval. This is the case
act, some of the covariates changed
udied, but were constant within the
each dosing interval, on average, 9
ion measurements were available,

allowing a reasonable empirical Bayes estimate of the individual parameters. For Drug B one of the covariates (*GLP*) changed markedly with each concentration measurement for each individual. To proceed with these data individual records were not subdivided by period because on average only 2.7 samples were available per individual. Instead, the time-averaged value of *GLP* (and other covariates) was used in Step 2.

The θ, Ω, and σ required for the empirical Bayes estimation were obtained by fitting the basic PK model to the modified data sets (with subdivisions when necessary) for drugs A–D. At this stage, parameters p_{ki} and p_{li}, $l \neq k$ were assumed to be uncorrelated, as previously used in the definition of Ω.

Initial Screening for Covariate Effects

Before fitting a GAM by stepwise addition/deletion, the individual PK parameter estimates are regressed independently on each covariate, which is equivalent to the method of Maitre *et al.* (2). This initial screening gives a first impression of the relative importance of several covariates (i.e., their ability to reduce the residual sum of squares) and of the shape of the relationships between covariates and PK parameters. The results of the initial screening for drug A are summarized in Table I, using the parameter clearance (*Cl*) as an example. The covariates were modeled linearly or as a natural cubic spline with two internal knots at the 33 and 66% quantiles (equivalent to 4 df). The significance (p) of the linear models was tested against the constant model using the F test. The *RSS* of the constant model was 12.15. The significance of the splines was tested against the linear models. For drug A the initial screening resolved that ($p < 0.05$) *AGE*, *WT*, *HT*, *DRG*, and *FED* significantly influenced *Cl*. The relationships between *Cl* and the continuous covariates (*AGE*, *WT*, *HT*) were best described by a linear model.

Table I. Change in Residual Sum of Squares (ΔRSS) After Independent Correlation of Each Covariate with the Clearance of Drug A

	Linear vs. constant		Spline vs. linear	
Covariate	ΔRSS	p	ΔRSS	p
WT	−3.89	< 0.01	−0.10	0.57
AGE	−3.88	< 0.01	−0.13	0.49
HT	−3.64	< 0.01	−0.21	0.33
FED	−0.61	0.03		
DRG	−0.58	0.03		
RACE	−0.46	0.06		
LOC	−0.07	0.45		
SEX	−0.03	0.60		

The use of a natural cubic spline to describe these relationships did not result in a significant improvement of the fit.

GAM Analysis

The GAM is then built using stepwise addition/deletion. At each substep of the GAM steps, each of the covariates is allowed to be left out of the model, to enter linearly, or to enter as a natural cubic spline with two internal knots at the 33 and 66% quantiles. For drug A the GAM analysis indicated that *Cl* is a linear function of *WT* and a function of the categorical covariates *DRG* and *RACE*, and *V* is a linear function of *WT* and a function of the categorical covariates *LOC* and *DRG*. Table II summarizes the path taken to these models and shows some of the other GAMs that are close to the best fit one. In the notation of the model equations, ~ means "is a function of" and + signs are to be interpreted figuratively and not literally. In each step the term that results in the largest decrease in the *AIC* is added to the model. The residual sum of squares (*RSS*) and change in *RSS* from the previous model are also shown. *WT* has the greatest influence on *Cl*, and it alone explains 32% of the variability in *Cl*, while all three factors (*WT*, *DRG*, and *RACE*) explain 38% of the variability. *HT* and *AGE*, which appeared to be important factors in the initial, one covariate at a time, screening ($p<0.01$), do not appear in the additive model. The reason for this is that these two covariates are highly correlated with *WT* ($r>0.940$)

Table II. Path Taken to the Final GAM for the Clearance (*Cl*) and Volume (*V*) Parameters of Drug A

Cl model[a]					*V* model[a]				
Step	Term	Δ*RSS*	*RSS*	*AIC*	Step	Term	Δ*RSS*	*RSS*	*AIC*
1	constant		12.16	234.3	1	constant		40.1	345.3
2	+ *WT*	−3.89	8.26	200.5	2	+ *HT*	−6.6	33.6	330.7
3	+ *DRG*	−0.36	7.90	198.2	3	+ *LOC*	−4.2	29.3	320.3
4	+ *RACE*	−0.32	7.58	196.3	4	+ *DRG*	−1.1	28.2	318.6

Summary of models close to the minimum model

Cl models		*V* models	
Equation[b]	*AIC*	Equation[b]	*AIC*
Cl ~ *WT* + *DRG* + *RACE*	196.3	*V* ~ *WT* + *LOG* + *DRG*	318.5
Cl ~ *WT* + *DRG* + *RACE* + *LOC*	197.3	*V* ~ *HT* + *LOC* + *DRG*	318.6
Cl ~ *WT* + *DRG* + *RACE* + *SEX*	197.3	*V* ~ *WT* + *LOC* + *DRG* + *DRG* + *SEX*	319.7
		V ~ *HT* + *LOC* + *DRG* + *SEX*	319.9
		V ~ *HT* + *LOC* + *DRG* + *FED*	320.0

[a] In each step the term that results in the largest decrease in the *AIC* is added (+) to the model.
[b] In the notation of the model equations ~ means "is a function of" and the + signs are to be interpreted figuratively and not literally.

: these relationships did not result

: addition/deletion. At each sub-
iates is allowed to be left out of
; a natural cubic spline with two
;. For drug A the GAM analysis
'and a function of the categorical
ar function of *WT* and a function
:G. Table II summarizes the path
the other GAMs that are close to
nodel equations, ~ means "is a
eted figuratively and not literally.
gest decrease in the *AIC* is added
(*RSS*) and change in *RSS* from
has the greatest influence on *Cl*,
lity in *Cl*, while all three factors
e variability. *HT* and *AGE*, which
initial, one covariate at a time,
: additive model. The reason for
y correlated with *WT* ($r > 0.940$)

arance (*Cl*) and Volume (*V*) Parameters

	V model[a]			
p	Term	Δ*RSS*	*RSS*	*AIC*
	constant		40.1	345.3
	+*HT*	−6.6	33.6	330.7
	+*LOC*	−4.2	29.3	320.3
	+*DRG*	−1.1	28.2	318.6

ne minimum model

V models	
Equation[b]	*AIC*
WT+*LOG*+*DRG*	318.5
HT+*LOC*+*DRG*	318.6
WT+*LOC*+*DRG*+*DRG*+*SEX*	319.7
HT+*LOC*+*DRG*+*SEX*	319.9
HT+*LOC*+*DRG*+*FED*	320.0

ase in the *AIC* is added (+) to the model.
a function of" and the + signs are to be

and therefore add nothing to it. Figure 1 shows a plot of all covariates found in the GAM for *Cl*, vs. a variable indicating their effect on *Cl* (see legend to Fig.). The spline for *WT* is displayed instead of the linear relationship found in the minimum model. Figure 1 clearly shows that the relationship between *Cl* and *WT* is predominantly linear. The nonlinearity is apparently caused by only one subject on the far right of the WT scale.

During the model search numerous models are visited. The GAM was derived using both a forward search starting with the NULL model and a backward search starting with the FULL model, which includes all covariates. In all, 71 different models were tested for *Cl*. For *Cl* the forward and backward search resulted in the same model. For *V* a small difference was found: *HT* appeared in the forward search, while *WT* appeared in the backward search. However, these two covariates are highly correlated which makes this understandable. The covariates *HT* (or *WT*) and *LOC* have the biggest effect on *V* and explain 27% of the variability. The addition of *DRG* results in a minor additional drop in *AIC*.

Final NONMEM Analysis

In the third and final step the population model is built using NONMEM on the basis of the GAM results of the previous step. At a first substep, only linear relationships between covariates and parameters are considered. In doing so, basically two approaches can be used: (i) Include in the initial NONMEM model all covariates that appear in the GAM with the lowest *AIC* value, and in other models close to it, or (ii) start with the minimum *AIC* model as the initial NONMEM model and test later whether the other covariates that appear in (Step 2) models close to the minimum *AIC* model significantly reduce −2 log likelihood (−2*LL*) obtained by the NONMEM analysis. (The difference in −2*LL* is asymptotically χ^2 distributed). The potential additional covariates can be tested according to their order of importance in Step 2. For Drug A approach (i) was used, because the total number of covariates was relatively small. The following models for *Cl* and *V* were used in the initial NONMEM model (the covariates are centered at their medians)

$$Cl = \theta_1 + \theta_2(WT - 15) + \theta_3 DRG + \theta_4 RACE + \theta_5 LOC + \theta_6 SEX \quad (6)$$

$$V = \theta_7 + \theta_8(WT - 15) + \theta_9 LOC + \theta_{10} DRG + \theta_{11} SEX + \theta_{12} FED \quad (7)$$

A covariance term between *Cl* and *V* was added to the model, because the plot of the empirical Bayes *Cl* and *V* estimates showed a strong correlation. The parameter estimates found in Step 2 were used as initial estimates for the NONMEM fit. In general, a good agreement was found between the parameter estimates in the GAM and NONMEM analyses. Including the

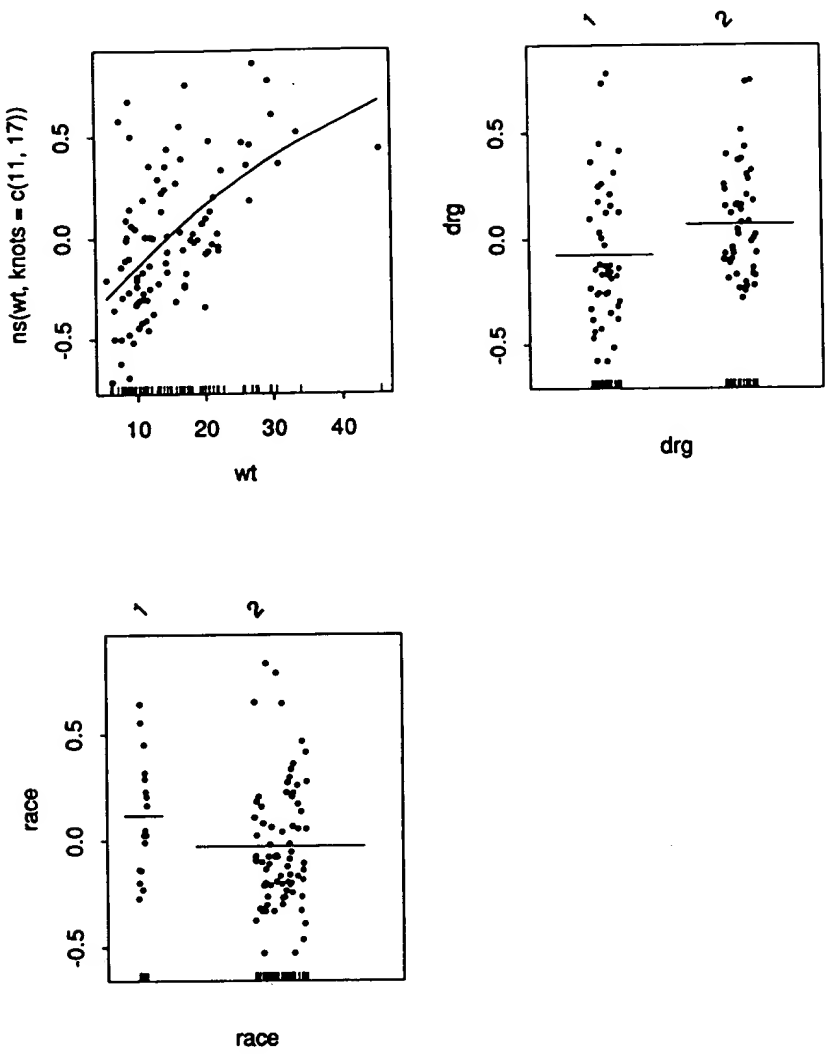
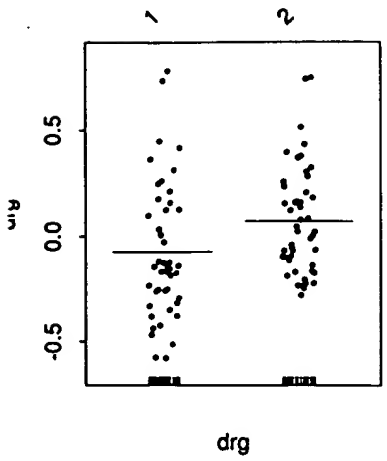


Fig. 1. Fit of the GAM for the *Cl* of drug A. Each plot is the contribution of a term to the additive predictor of *Cl*. The ordinate label is the expression used to specify the contribution of the covariate to the model formula in the S-plus language. The ordinate value can be thought of as the (partial) effect of the covariate on *Cl*. Each curve has been centered to have an average of zero. The points plotted are the partial residuals: That is, the individual *Cl* values minus the prediction based on all other covariates. The same scale is used for the y axis in all plots so that the relative importance of the covariates can be compared. The two dose level (*DRG*) groups were labeled 1 = 5 mg/kg and 2 = 10 mg/kg, the race (*RACE*) groups were 1 = caucasian, 2 = black. For these dichotomous covariates, points are randomly shifted by small amounts on the abscissa to display them better.



h plot is the contribution of a term to the
xpression used to specify the contribution
lus language. The ordinate value can be
on *Cl*. Each curve has been centered to
e partial residuals: That is, the individual
covariates. The same scale is used for the
of the covariates can be compared. The
g/kg and 2 = 10 mg/kg, the race (*RACE*)
hotomous covariates, points are randomly
ay them better.

various covariates in the model results in a highly significant drop of 238 in $-2LL$ when compared to the basic (Step 1) PK model. The initial [Eqs. (6) and (7)] model was further refined by deleting certain parameters from the model according to the following criteria: (i) Only a small change in $-2LL$ when the parameter is left out, (ii) ± 1 SE of the parameter estimate includes zero, and (iii) the magnitude of the covariate coefficient is small (i.e., the covariate has little effect on the PK parameter). This resulted in the following models for Cl and V in the final NONMEM model

$$Cl = \theta_1 + \theta_2(WT - 15) + \theta_3 DRG + \theta_4 RACE \quad (8)$$

$$V = \theta_5 + \theta_6(WT - 15) + \theta_7 LOC + \theta_8 DRG + \theta_9 SEX \quad (9)$$

The influence of *LOC* and *SEX* on *Cl* and the influence of *FED* on *V* was not significant. The θs of these covariates were poorly determined with a standard error larger than the parameter estimate. Fixing the θs of these covariates to zero resulted in an increase of only 3.4 in $-2LL$. Thus, the parameters that could be deleted from the initial model appeared to be the ones that were the least important according to the covariate selection in Step 2. Table III summarizes the change in $-2LL$ when each of the θs of the covariates appearing in the reduced (and final) NONMEM model is set to zero, showing the significance of each of the parameters (a change > 6.6 in $-2LL$ is significant with $p < 0.01$). For *Cl*, *WT* seems to be the most important covariate in the final population model. This is in accordance with the GAM analysis: *WT* resulted in a large decrease in *AIC*, while addition of the other two covariates resulted in a relatively small additional decrease. For *V*, *WT* and *LOC* appear to have the greatest effect, which is also in accordance with the results found in the GAM analysis. For *V* an additional covariate, *SEX*, that appeared in a GAM close to the minimum model, was significant. Of the four covariates not found in the minimum *AIC* GAM but in models close to it (Table II), only *SEX* significantly influenced the fit. Including the final set of covariates in the NONMEM model resulted in a decrease of the interindividual variability from 46 to

Table III. Change in $-2LL$ When Each of the θs of the Covariates Appearing in the Final NONMEM Model for Drug A is Set to Zero

Cl model		*V* model	
Covariate	$-2\Delta LL$ (vs. $\theta = 0$)	Covariate	$-2\Delta LL$ (vs. $\theta = 0$)
WT	115	*WT*	77.5
DRG	16.0	*LOC*	50.2
RACE	9.9	*DRG*	8.3
		SEX	14.4

29% for Cl and from 46 to 28% for V when compared to the basic PK model without covariates.

Validation of Final NONMEM Model

As a first check on the NONMEM analysis, the relationship between WT and Cl and WT and V was modeled by a natural cubic spline (see Appendix) with two internal knots at the 33 and 66% quantiles (4 df). This nonparametric representation did not result in a significant change in $-2LL$, confirming the results found in the GAM analysis that these relationships are predominantly linear.

Finally, to check whether our final model was really the minimum model, all other covariates were included in the model one by one, but none significantly influenced $-2LL$. Another, quicker approach to test whether the final model is the minimum model may be the following. Calculate the $\exp(\eta_{kj})$ (the unexplained residual interindividual variability) as empirical Bayes estimates using the final NONMEM model. Then, regress the $\exp(\eta_{kj})$s on the covariates that are not in the final NONMEM model. This was done for drug A and, as expected, none of the remaining covariates correlated significantly with the final model-based empirical Bayes estimates. We propose this as a final step in the analysis to confirm whether an appropriate final model has been reached.

Comparison of GAM and NONMEM Analyses

The GAM analysis apparently provided both the important covariates for the final NONMEM model and the shapes of the relationships between covariates and PK parameters for example A. An important question regarding the suggested approach, however, is whether the GAM analysis always suggests all the relevant covariates. Although we cannot know that for sure, experience with drugs B–D suggests that it is so. Table IV compares the covariates that were selected by the GAM approach and the covariates that were significant in the population model derived using NONMEM. For all drugs the GAM with the lowest AIC included all the important covariates or they were included in GAM models close to the minimum model. This latter case happened only once: Drug B, covariate RF. Table IV shows the results only for the Cl parameter of the four drugs, but the same findings held for V. The GAM fit also identified the relative importance of the covariates. In general, the covariates that resulted in the biggest decrease of the AIC in the GAM step were also the most important covariates in the NONMEM model (judged by change in $-2LL$). Conversely, the covariates that were of borderline significance in the GAM fit were of minor significance in the final NONMEM model or could be deleted from the model.

:n compared to the basic PK model

analysis, the relationship between
led by a natural cubic spline (see
: 33 and 66% quantiles (4 df). This
ult in a significant change in $-2LL$,
M analysis that these relationships

:l model was really the minimum
in the model one by one, but none
quicker approach to test whether
ıay be the following. Calculate the
individual variability) as empirical
MEM model. Then, regress the
n the final NONMEM model. This
none of the remaining covariates
:el-based empirical Bayes estimates.
ılysis to confirm whether an appro-

ıalyses

ided both the important covariates
hapes of the relationships between
e A. An important question regard-
whether the GAM analysis always
ugh we cannot know that for sure,
:t it is so. Table IV compares the
1 approach and the covariates that
derived using NONMEM. For all
:luded all the important covariates
:lose to the minimum model. This
covariate *RF*. Table IV shows the
four drugs, but the same findings
e relative importance of the covari-
lted in the biggest decrease of the
important covariates in the NON-
L). Conversely, the covariates that
M fit were of minor significance in
deleted from the model.

Table IV. Covariates Found by the GAM and NONMEM Analysis to Significantly Influence *Cl*

Analysis	Covariates[a]							
				Drug A				
GAM	*WT*	*DRG*	*RACE*					
NONMEM	*WT*	*DRG*	*RACE*					
				Drug B				
GAM	*GLP*	*WT*	*RACE*	*ET*				
NONMEM	*GLP*	*WT*	*RACE*	*ET*	*RF*[b]			
				Drug C				
GAM	*HT*	*SEX*	*RACE*	*TOB*	*AGE*	*HCTZ*	*SECR*	*WT*
NONMEM	*HT*		*RACE*		*AGE*	*HCTZ*		
				Drug D				
GAM	*CLCR*	*BIL*	*WT*	*AGE*	*AP*	*CEN*	*SEX*	*SPB*
NONMEM	*CLCR*	*BIL*	*WT*	*AGE*		*CEN*		*SPB*

[a]The covariates are listed in their order of appearance in the GAM.
[b]Found in a GAM model close to the minimum *AIC* model.

Nonlinear Covariate Relationships

Another important question is whether the GAM step identifies important nonlinearities in the relationships between covariates and PK parameters. In this respect, a nice example is provided by drug C. In the GAM with the lowest *AIC*, *HT* and *AGE* were two continuous covariates that were of importance. Figure 2 shows the contribution of these two covariates to the GAM (the effects of the other covariates are not shown). The relationships were characterized by a natural cubic spline with two internal knots

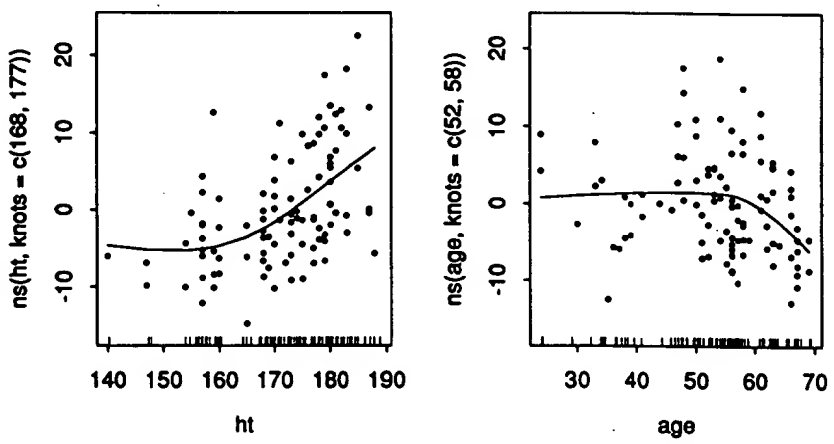


Fig. 2. Plot of the contribution of *HT* and *AGE* to the GAM for *Cl* of drug C (other relevant covariates are not shown). The relationships are described by a natural cubic spline with two internal knots at the 33 and 66% quantiles. See legend to Fig. 1 for further details.

at the 33 and 66% quantiles. The shapes found in the GAM suggest that some nonlinear representation may be appropriate for the NONMEM model. Subsequently, in the NONMEM fit a natural cubic spline (with two internal knots at the 33 and 66% quantiles, constrained to monotonicity) was used to describe the relationship between *Cl* and *HT* and *Cl* and *AGE*. Figure 3 shows the shapes of the spline representations found in NONMEM. Note the similarity between the splines found in the GAM analysis and those from the NONMEM fit. Using the splines in the NONMEM fit resulted in a drop of 30.0 points in $-2LL$ (4 additional parameters) when compared to the fit using a linear relationship. The particular shape of the splines suggest that *HT* only has an effect on *Cl* when the subjects are taller than 160 cm. Age only appears to influence *Cl* when the subjects are older than 60. Therefore the natural cubic splines were replaced by simpler (linear spline) models: *Cl* was assumed to be linearly related to *HT* for subjects taller than 160, and constant for smaller subjects; *Cl* was linearly related to *AGE* for subjects older than 60 years and constant for younger subjects. This simplified model resulted in about the same $-2LL$ as the original spline fit, but used fewer parameters. The nonlinearities in *HT* and *AGE* were both statistically significant. Replacing the nonlinear model for *AGE* by a linear model resulted in an increase of $-2LL$ of 18.8. Replacing the model for *HT* by a linear model increased the $-2LL$ by 7.4 points.

For drug D the GAM analysis selected a nonlinear relationship between *Cl* and *WT*. A similar relationship was found by NONMEM, resulting in a

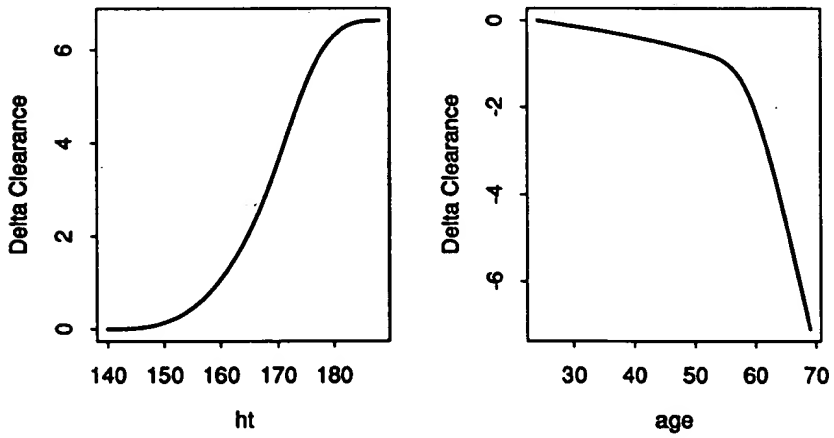
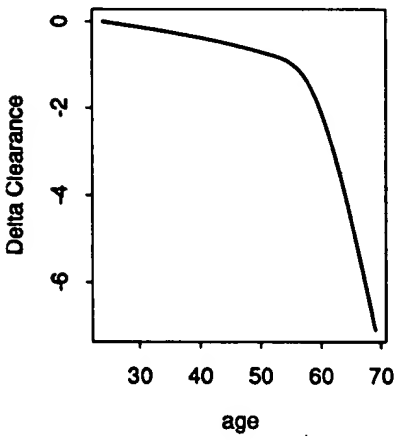


Fig. 3. Spline (natural cubic spline with two internal knots at the 33 and 66% quantiles, constrained to monotonicity) representation of the relationship between *Cl* and *HT*, and *Cl* and *AGE* for drug C found in NONMEM fit.

found in the GAM suggest that some
opriate for the NONMEM model.
tural cubic spline (with two internal
strained to monotonicity) was used
and *HT* and *Cl* and *AGE*. Figure 3
itations found in NONMEM. Note
in the GAM analysis and those from
the NONMEM fit resulted in a drop
rameters) when compared to the fit
lar shape of the splines suggest that
subjects are taller than 160 cm. Age
ubjects are older than 60. Therefore
by simpler (linear spline) models: *Cl*
HT for subjects taller than 160, and
inearly related to *AGE* for subjects
nger subjects. This simplified model
e original spline fit, but used fewer
nd *AGE* were both statistically sig-
for *AGE* by a linear model resulted
cing the model for *HT* by a linear
s.
ted a nonlinear relationship between
found by NONMEM, resulting in a



ternal knots at the 33 and 66% quantiles,
the relationship between *Cl* and *HT*, and
t.

significantly better fit than the linear model. The correspondence between GAM and NONMEM analyses was not always perfect, however. For drug B, a nonlinear relationship between *Cl* and *GLP* was preferred in the final NONMEM model but was not detected by the GAM analysis. This was probably caused by the previously noted fact that the *GLP* changed markedly from one concentration measurement to the next within individuals, but because too few data were available to provide an empirical Bayes estimate of *Cl* for each value of *GLP*, the time averaged value of *GLP* for each individual was correlated with the individual *Cl* estimates in the GAM step. This undoubtedly obscured the predominantly intraindividual correlation between the two, which NONMEM eventually discovered.

DISCUSSION

NONMEM is a widely used program to determine the PK and PD parameters of a drug in a patient population. One of the advantages of developing a population PK-PD model is that the influence of certain factors, such as (patho)physiological conditions, concomitant therapy, etc., on the drug's pharmacokinetics and pharmacodynamics can be included in the population model. This allows identification of possibly causal effects that can later be confirmed, the design of patient-specific dosage regimens, and optimization of drug dosage adjustment using Bayesian approaches.

Finding a population model that adequately describes the data can be a complicated task. Recently, Maitre *et al.* (2) proposed a method that forms the basis of the method presented here. They proposed to look at scatterplots of empirical Bayes estimates of the model parameters versus the covariates in order to detect which covariates to use in the population model. The problem with this approach is that it is valid only if the covariates act independently on the PK-PD parameters. In this paper we elaborate on Maitre *et al.*'s (2) approach and propose a more formal analysis of the empirical Bayes estimates in which we regress the individual empirical Bayes estimates on the candidate covariates using a generalized additive model (GAM). An important feature of the GAM is that it provides a functional representation of the relationship between PK-PD parameters and covariates. Thus using the GAM analysis, a model that can describe the relationship between PK-PD parameters and the covariates is built outside NONMEM. At a later stage, NONMEM is used to optimize and finalize the tentative model thus provided. The great advantage of this elaboration is speed. The total time required to build the population model is reduced dramatically because the number of NONMEM runs is reduced. The elucidation of the GAM that best describes the relationship between the individual estimates of the model parameters and the covariates is done quickly relative

to the series of NONMEM runs required to perform the same step. Furthermore, the data analysis tools used in the GAM step give a nice graphical representation of the relationship between model parameters and covariates, something heretofore lacking in the NONMEM software.

The examples presented in this paper have shown that the derived GAM is a very good initial model for the NONMEM analysis. Apart from selecting the most important covariates the GAM approach also indicates important nonlinearities in the relationships between covariates and PK parameters. An important feature of the GAM step is building a model that includes the contribution of several covariates. This model provides a better starting point than the initial, one covariate at a time, screening (which is equivalent to the Maitre *et al.* suggestion). The initial, one variable at a time, screening tends to select numerous covariates that do not appear to be of importance in the final NONMEM model, usually because they carry the same information as one or a combination of the other covariates. What is even more important, the initial screening misses some of the covariates that appear to be of importance in the final NONMEM model. The reason here may be that their effect is overshadowed by some other covariate or that correlation among covariates masks the relationship. A nice example is provided by drug D. Of the 16 reported covariates, 13 were found to significantly affect *Cl* according to the one-at-a-time initial screening ($p<0.05$, F test). The GAM model narrowed this down to 8 covariates of which 6 were finally selected by the NONMEM analysis. One of the covariates finally selected for the NONMEM model (*SPB*) was not selected by the initial screening but did appear in the GAM. Another example is drug A. The one-at-a-time initial screening selected 6 covariates to significantly affect *Cl* (Table I), of which 2 appeared in the final NONMEM model (Table III). However, one covariate (*RACE*), which is significant in the final NONMEM model, was not selected during the initial screening. Conversely, the GAM analysis selected the same 3 covariates as appeared in the final NONMEM model (Tables II and III). Similar results were found for the other two drugs.

A disadvantage of the approach is that its success depends on the quality of the individual empirical Bayes parameter estimates. These estimates tend to be "shrunken" towards the population mean, especially when few data are available from an individual (or period within an individual). This bias of the empirical Bayes estimates complicates the detection of relevant covariates, and perhaps even more, it complicates the elucidation of the shape of the relationship between the model parameters and the covariates. However, if a covariate is found to be of importance in the GAM analysis, it has a greater chance of being clinically relevant. The same probably holds for shape effects. If an important nonlinearity exists we surmise that it will either show up in the GAM or a linear representation may be an appropriate

d to perform the same step. Further-
he GAM step give a nice graphical
en model parameters and covariates,
)NMEM software.
er have shown that the derived GAM
IMEM analysis. Apart from selecting
M approach also indicates important
/een covariates and PK parameters.
is building a model that includes the
is model provides a better starting
ı time, screening (which is equivalent
tial, one variable at a time, screening
ıt do not appear to be of importance
)ecause they carry the same informa-
)ther covariates. What is even more
.ome of the covariates that appear to
:M model. The reason here may be
ne other covariate or that correlation
hip. A nice example is provided by
13 were found to significantly affect
ial screening ($p < 0.05$, F test). The
3 covariates of which 6 were finally
)ne of the covariates finally selected
not selected by the initial screening
xample is drug A. The one-at-a-time
ɔ significantly affect *Cl* (Table I), of
:M model (Table III). However, one
. in the final NONMEM model, was
ng. Conversely, the GAM analysis
:ared in the final NONMEM model
e found for the other two drugs.
:hat its success depends on the quality
neter estimates. These estimates tend
:ion mean, especially when few data
·riod within an individual). This bias
:ates the detection of relevant covari-
.cates the elucidation of the shape of
ımeters and the covariates. However,
tance in the GAM analysis, it has a
vant. The same probably holds for
ity exists we surmise that it will either
)resentation may be an appropriate

approximation. So far we have observed that the GAM analysis is inclusive. For the four drugs a total of 27 covariates were found to be of significance in the NONMEM models. Of these covariates 25 were found in the GAM with the lowest *AIC* value for the drugs. The remaining 2 were found in models close to the minimum *AIC* model. Thus, apparently the shrinkage is not a major problem in detecting important covariates. As mentioned in the Data section, an additional problem is that interactions between covariates are not included in the GAM analysis. Recently, two new approaches to covariate and shape selection: PI (9) and MARS (10) have become available. These approaches automatically search for combinations of covariates showing important interactions and determine their functional representation using splines. We are currently investigating the use of these approaches at Stage 2 and the implementation of the derived models at Stage 3.

A final note of caution: In model building subjective judgment plays a much larger role than in other more formal pursuits. The final model is not the only, or true one; only a useful one, and although our experience is encouraging, further work with the approach we suggest is required to determine its true utility.

ACKNOWLEDGMENTS

We thank Eli Lilly, Indianapolis, Indiana; Bristol Myers, Hillside, New Jersey; Rhone Poulenc Rorer, Antony, France, and Thomas M. Ludden for letting us use their data.

APPENDIX

A spline function, ($sp(x)$, where x is a predictor variable), is characterized by a strictly increasing sequence of real values ($b_1 < \cdots < b_n$) called knots, and by its order k. In the interval (b_i, b_{i+1}) the spline is a polynomial of order $k-1$, and across each b_i, derivatives of order 0 to $k-2$ are continuous. All spline calculations in NONMEM were made using the package (B-spline) based on PPPACK (11). B-spline is available from D. Verotta, Box 0626, University of California-San Francisco, San Francisco, CA 94143, upon request (email david@c255.ucsf.edu or david@ucsfc255.bitnet).

REFERENCES

1. S. L. Beal and L. B. Sheiner. Estimating population kinetics. *CRC Crit. Rev. Biomed. Eng.* 8:195–222 (1982).
2. P. O. Maitre, M. Bührer, D. Thomson, and D. R. Stanski. A three-step approach combining Bayesian regression and NONMEM population analysis: Application to midazolam. *J. Pharmacokin. Biopharm.* 19:377–384 (1991).

3. L. B. Sheiner, S. Beal, B. Rosenberg, and V. V. Marathe. Forecasting individual pharmacokinetics. *Clin. Pharmacol. Ther.* **26**:294–305 (1979).
4. L. B. Sheiner and S. Beal. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. *J. Pharm. Sci.* **71**:1344–1348 (1982).
5. T. J. Hastie and R. J. Tibshirani. *Generalized Additive Models*, Chapman and Hall, New York, 1990.
6. R. E. Bellman. *Adaptive Control Processes*, Princeton University Press, Princeton, NJ, 1961.
7. T. J. Hastie. Generalized additive models. In J. M. Chamber and T. J. Hastie (eds.), *Statistical Models*, Wadsworth and Brooks, Pacific Grove, CA, 1991, pp. 249–304.
8. H. Akaike. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**:716–723 (1974).
9. L. Breiman. The PI method for estimating multivariate functions from noisy data. *Technometrics* **33**:125–160 (1991).
10. J. H. Friedman. Multivariate adaptive regression splines. *Ann. Statist.* **19**:1–141 (1991).
11. C. A. DeBoor. *A Practical Guide to Splines*, Springer-Verlag, New York, 1978.